22.7 A Single-Cycle-Access 128-Entry Fully Associative TLB for Multi-Core Multi-Threaded Server-on-a-Chip

Shashank Shastry, Ajay Bhatia, Sagar Reddy

Sun Microsystems, Sunnyvale, CA

The translation lookaside buffer (TLB) is a timing critical component of the virtual memory management unit (MMU) of high-performance microprocessors. This TLB has a single-cycle access with embedded way selection and cache-hit detection logic, a dual-storage CAM cell, a modified dual-matchline true hit detect circuitry and a dual-port data RAM, for faster virtual address (VA) to physical address (PA) translation [1]. The TLB is shared by the eight threads of each of the eight SPARC™ cores of the Niagara2 microprocessor [4].

This 128-entry fully associative TLB accepts a 48b VA and two 13b contexts and translates to a 40b PA. Figure 22.7.1 shows the basic architecture of the TLB. The TLB is implemented as a CAM and RAM array. The CAM array stores the virtual or real address tag while the 1read/1write (1R1W) register-file-based RAM array stores the physical page number along with page size bits and other page attribute bits. The TLB compares the partition identifier, virtual page number, real bit and context with each entry of the TLB. Any single entry matched is qualified with the valid bit to generate the PA. The corresponding used and valid bit of a matched entry is updated during a compare or demap operation. The replacement policy used is a pseudo-least-recently-used (pseudo-LRU) algorithm, implemented using a 128b priority encoder that receives 128 valid bits and 128 used bits as input. The replacement algorithm replaces the first entry that is not used or not valid. When all the used bit entries are set, the TLB resets all the entries of the used bit array. The TLB supports multiple page sizes ranging from 8KB to 256MB. The PA from the data RAM is compared with four way-addresses from the cache tags to determine the cache hit using a 30b comparator. The TLB also checks for multiple CAM hits on each access. In bypass mode, the virtual address is forwarded to the PA. A 38b even parity of the translated data is generated.

Figure 22.7.2 shows the circuit architecture of the CAM array. The CAM match circuitry is organized into local and global matchlines. There are 3 CAM cells on each miss-discharge local matchline and 1 CAM cell on each hit-discharge local matchline. All local match lines are precharged high while the clock is low. A miss-discharge matchline stays precharged on a match and evaluates low on a miss (XOR functionality) while a hit-discharge matchline stays high on a miss and asserts low on a hit (XNOR functionality). Monotonically rising compare data is XORed or XNORed with the stored complementary data on rising edge of the clock. The dual matchline circuitry consists of a global missline and a global hitline to detect a true hit.

Figure 22.7.3 shows the schematic of the TLB CAM cell and the first stage local matchline. The CAM cell has dual storage, which consists of two 6T SRAM cells, enabling the masking function required for variable page size and context array. The unmasked bits are stored as complementary data in each cell of the dual-storage CAM cell. The bits of the VA or context that must be masked due to page size or the real bit being set, are stored as logic 0 rather than the true and the complement of the bit value. The masked bits do not participate in a match since they cannot discharge the local matchline. The layout of the TLB CAM cell is implemented in L-shaped pairs to save area and a single polysilicon gate orientation is maintained to minimize critical dimension variations.

Figure 22.7.4 shows the schematic of the modified dual-matchline circuitry and Fig. 22.7.5 shows the logical states of the various matchlines. The global matchline circuit performs a logical AND of all local match results. The passgate output *ml_sa* is precharged high during the low phase of clock, hence the output of dual matchline *ml_out* is always low (default miss state). The NMOS pull down transistor stack of the global hitline *ml_hit* receives the XNOR (shown as NOR) result of first bit comparison while the global missline *ml_miss*, a wired-OR, receives the XOR (shown as NAND) result of the remaining bits comparison. In the event of a hit, *ml_hit* discharges while *ml_miss* stays precharged, which allows *ml_sa* to evaluate low, causing *ml_out* to assert high and indicating a true hit. In the event of a single- or multi-bit miss, both *ml_hit* and *ml_miss* discharge. The pulldown stack on *ml_hit* ensures a slower slew rate as compared to *ml_miss*. This results in zero or negative gate overdrive of NMOS passgate transistor and hence, *ml_out* stays low. To match the slew rate of *ml_hit* and *ml_miss*, an exact replica dummy load of all NMOS devices is connected on *ml_hit*. The dual-matchline circuit tracks process variations due to NMOS-only implementation of both the matchlines. The layout of the matchlines is metal- and load-matched and kept local to reduce capacitance mismatch and ensure common-mode noise rejection. The passgate has a higher threshold voltage and higher-than-minimum channel length to improve yield. A nested Monte Carlo analysis with global and local mismatch of devices and worst-case wire variations in the presence of varying skew between *ml_hit* and *ml_miss* was simulated. The results show a skew tolerance of 6 gate delays between *ml_hit* and *ml_miss.*

The *ml_out* signal is also used as a wordline to the RAM array by forcing a hit on the global hitline as described in Fig. 22.7.5. The conventional requirement of having synchronous control circuitry between the CAM and RAM is eliminated. Multiple data-RAM wordlines can be ON without destroying the contents of the RAM array. This allows for easier silicon debug and multi-match circuit implementation. The 128-entry data array is organized as two 64-entry hierarchical arrays. Each 64-entry data RAM array consists of 8 sub-arrays with 8 entries each. A full-swing local and global read bitline scheme is used to achieve speed and allow for low voltage scaling. The data-RAM bitcell is an 8T 1R1W SRAM cell. This cell is chosen instead of a conventional 6T SRAM to avoid reliability and short circuit problems in the case of a multiple entry match. The dual matchline and register-file RAM implementation allows for flow-through access of the TLB path and does not have skew or clock synchronization penalties or the requirement to wait for clock edges to initiate operations [2,3].

The TLB has an associated BIST on the Niagara2 microprocessor to test read/write operations from CAM array, RAM array, valid-bit/used-bit arrays. It also tests various compare/demap operations, multi-hit and replacement algorithm. Both 6T bitcells of the TLB CAM cell are accessed with a single sense-amplifier and a programmable recirculating flip-flop to select between the two. The skew between *ml_hit* and *ml_miss* is controlled with a programmable delay chain using recirculating scannable flip-flops. All input and output flip-flops are scannable. The TLB is also testable by a slow scan macrotest mode to improve test coverage.

A 128-entry data TLB is implemented in a 65nm, triple-threshold voltage, 1.1V CMOS process. A 64-entry instruction TLB was similarly implemented. BIST and macrotest were successfully run. Figure 22.7.6 shows the Niagara2 microprocessor micrograph.

References:

[1] Y.-W. Ho, I. Bhasin, T. Chiu, et al., "A Process-Portable 64b Embedded Microprocessor with Graphics Extensions and A 3.6GB/s Interface," *ISSCC Dig. Tech. Papers*, pp. 234–235, Feb., 2001.

[2] R.A. Heald and J.C.Holst, "A 6-ns Cycle 256-kb Cache Memory and Memory Management Unit," *IEEE J. Solid-State Circuits*, vol. 28, no. 11, pp. 1078–1083, Nov., 1993.

[3] L.R Tamura, T.-S. Yang, D.E.-Wingard, et al., "A 4-ns BiCMOS Translation-Lookaside Buffer," *ISSCC Dig. Tech Papers*, pp. 68–69, Feb., 1990

[4] G. Grohoski, "Niagara2: A Highly-Threaded Server-on-A-Chip," *Hot Chips Proceedings*, Aug., 2006.

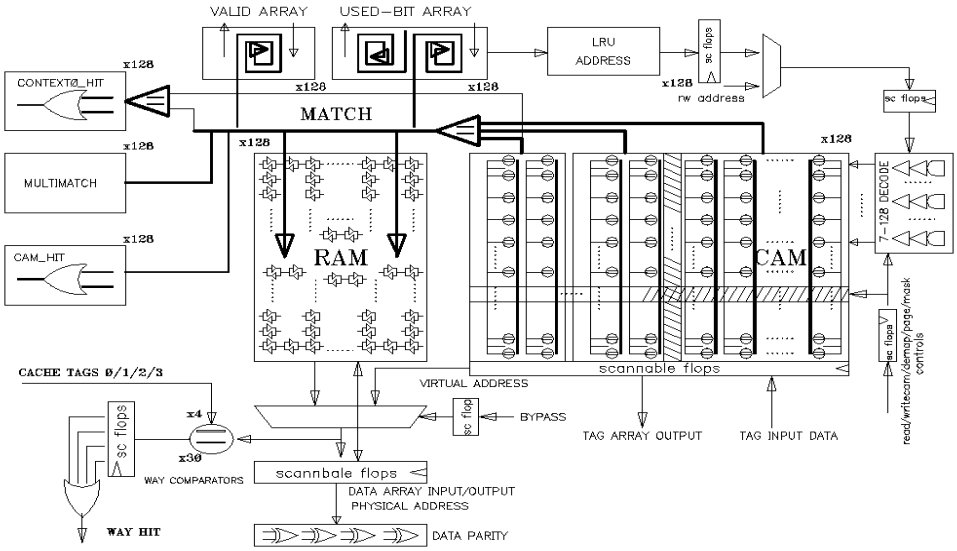

Figure 22.7.1: TLB block diagram.

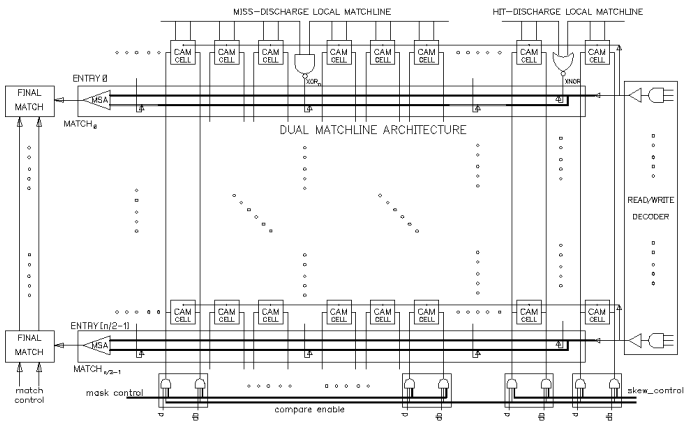

Figure 22.7.2: CAM array circuit architecture.

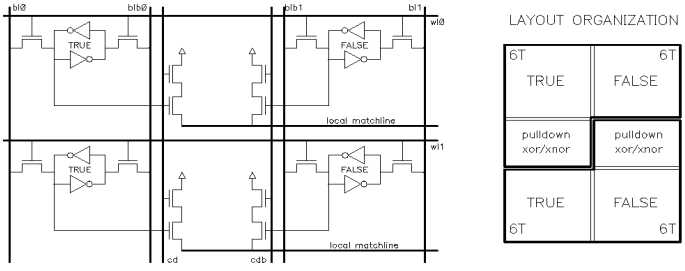

Figure 22.7.3: TLB CAM cell.

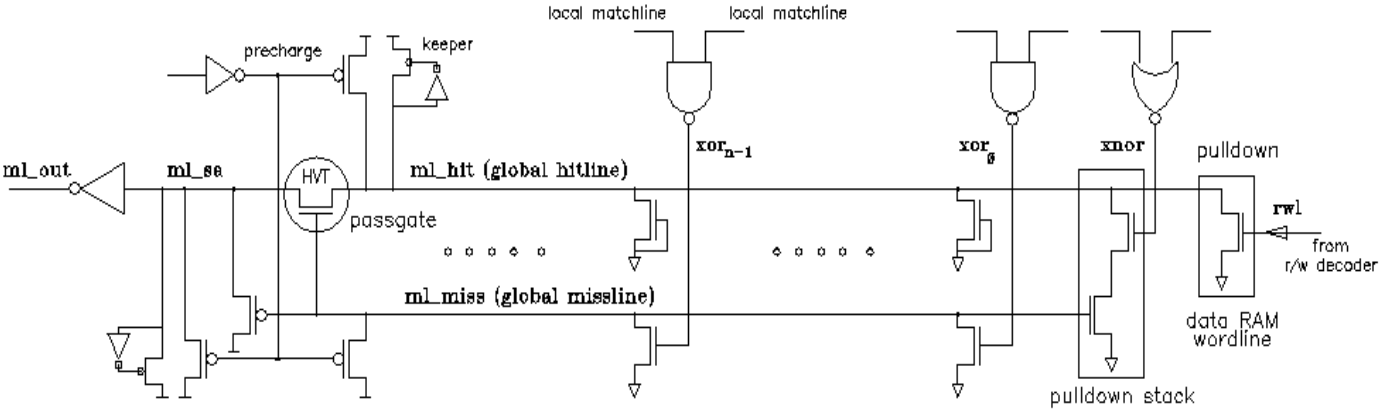

Figure 22.7.4: TLB dual matchline.

Continued on Page 612

case	OPERATION	rwl	xnor	xor0	ml_hit	ml_miss	ml_sa
1	compare	0	1	0	1 -> 0	1 -> 1	1 -> 0
2	compare	0	0	0	1 -> 1	1 -> 1	1 -> 1
3	compare	0	0	1	1 -> 1	1 -> 0	1 -> 1
4	compare	0	1	1	1 -> 0	1 -> 0	1 -> 1
5	read/write	1	0	0	1 -> 0	1 -> 1	1 -> 0

Figure 22.7.5: TLB dual matchline switching table.

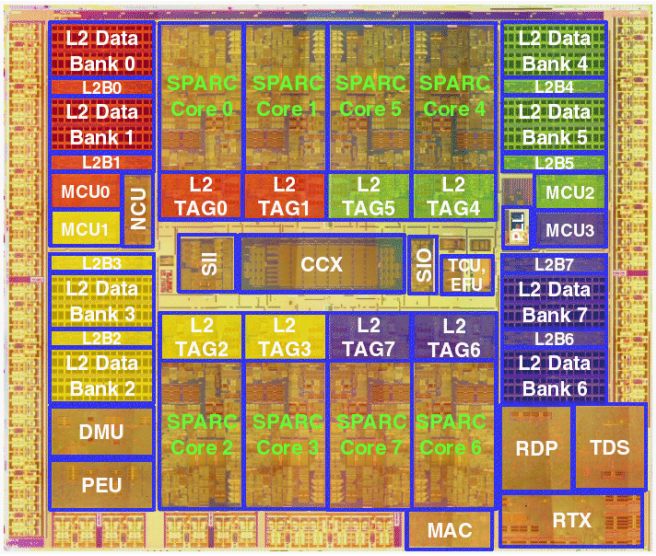

Figure 22.7.6: Niagara2 SPARC microprocessor micrograph.

